Notes and Documents

The Population of the United States, 1790: A Symposium

[Editor's Note: The elements of this symposium came together through a combination of authorial initiative and editorial design. Mr. Purvis's article was approved for print in mid-November 1982. By a happy coincidence, Mr. Akenson's appeared in our office only a few days later. After each author had agreed to exchange comments, their articles were sent for further comment to Mr. and Mrs. McDonald. Pleased with the package, we do not intend to prolong the discussion in these pages, save for such responses as the participants may make to letters from our readers, which we welcome.]

The European Ancestry of the United States Population, 1790

Thomas L. Purvis

FORREST and Ellen Shapiro McDonald have been at the center of controversy ever since they began reexamining the national origins of white Americans at the taking of the first census and reinterpreting southern history with their Celtic Thesis.[1] Lost in the ensuing debate has been any recognition that they deserve much credit for alerting historians to the possibility of error in the most commonly cited figures on the European descent of the colonial white stock and for demonstrating

Mr. Purvis is a member of the Department of History at the University of Georgia. He wishes to thank Ned Landsman for his generosity in making his unpublished research available and for commenting on the author's work.

[1] Forrest McDonald and Ellen Shapiro McDonald, "The Ethnic Origins of the American People, 1790," *William and Mary Quarterly*, 3d Ser., XXXVII (1980), 179-199.

how to improve the precision of such statistics. It is therefore extremely regrettable that the McDonalds' recent reassessment of the work of Howard F. Barker and Marcus L. Hansen on the distribution of national groups in 1790 included only British Americans, for the research of Barker and of Hansen commonly serves as a starting point for ethnic history in the United States.[2] The McDonalds revised none of their predecessors' estimates of Germans, Dutch, Swedes, and French, nor did they update any figures for five of the six states whose census rolls for 1790 are missing: New Jersey, Delaware, Georgia, Kentucky, and Tennessee.

Barker and Hansen's conclusions on the size of Continental nationalities in the American population were even more questionable than their findings for Englishmen and Celts. Barker's examination of Germans also suffered from shortcomings similar to those the McDonalds have criticized regarding his treatment of British stocks. His figures on the distribution of ethnic groups in New Jersey, Delaware, Georgia, Kentucky, and Tennessee were even more suspect. Because their 1790 census rolls had burned, he made estimates for these five by extrapolating from the mix of national stocks in neighboring states, a very inexact procedure. Because Hansen believed the minor stocks were too small to justify using arithmetic computations for projections of their size, his research on the Dutch, Swedes, and French was neither quantitative nor systematic. The accuracy of Hansen's work consequently is open to some doubt.

The purpose of this note is to continue the work begun by the McDonalds in developing more trustworthy approximations of national descent. It will devise new lists of distinctive surnames for determining the proportions of Continental European elements in the American population, similar to the McDonalds' improved testing procedures for Celts. It will also exploit alternative sources for the federal population schedules that were not available to Barker and Hansen, in order to supply more reliable statistics for the six states lacking census manuscripts from 1790. It will then be possible to make the first comprehensive reevaluation of how significantly emigration from different European nations contributed to early American population growth since Barker and Hansen's work appeared in 1932.

Barker's basic premise for investigating the size of the largest national groups was valid, as even his critics agree. Distinctive surnames known to be borne by a certain percentage of a European group can be used to calculate an arithmetic coefficient, which will be sufficiently accurate to compute the proportion of people belonging to that nationality within a given state. The number of individuals holding those same surnames, multiplied by the appropriate numerical constant, will equal the approxi-

[2] American Council of Learned Societies, "Report of Committee on Linguistic and National Stocks in the Population of the United States," American Historical Association, *Annual Report for the Year 1931*, I (Washington, D.C., 1932), 107-441, hereafter cited as "National Stocks."

mate size of the group within that state. A more direct methodology would certainly be preferable, but the fact is that this is the only way to develop statistical measurements of ethnicity within a population of mixed European descent when no other information is available.

Barker's underlying procedures were sound, but, as often happens in pioneering studies, his implementation suffered from several deficiencies. The McDonalds contributed two basic suggestions for revising Barker's findings. They noted that the separate calculation of English surnames, along with those of all other groups, introduced an unnecessary margin of error because the resulting figures always failed to equal 100 percent. Barker's solution for this problem was altogether unsatisfactory; he either relied on arbitrary assumptions to reduce the discrepancies or he left substantial portions of many states' populations unassigned, sometimes more than one-fifth of all whites. The simplest way to eliminate the potential for bias and inaccuracy inherent in this predicament is to subtract the cumulative percentage of all minority stocks from 100 percent, on the presumption that the remainder was overwhelmingly English. Equally important was the McDonalds' demonstration that the number of distinctive surnames used by Barker was so small that it unnecessarily increased the risk of statistical skewing. A larger sample of names would represent a greater percentage of the European population under investigation and consequently would produce a smaller, more precise constant with which to approximate that nationality's size.

Completing the evaluation of Barker and Hansen's findings begun by the McDonalds requires, first, that lists of characteristic family names and corresponding coefficients be developed for the continental European nationalities. An immediate problem arises because data similar to those compiled on naming patterns in Great Britain do not exist for Germany, the Netherlands, Sweden, or France. Information of this type would be of doubtful usefulness in any regard, since emigration from those countries was usually small and may have derived from regions where the frequency of common names varied markedly from the general population. Moreover, groups such as the Dutch and Swedes were not truly homogeneous in European descent; they were actually amalgams of several different national stocks that had adopted a common language and religion before the imposition of English authority.[3] The most accurate way of determining nomenclatural patterns among the non-British groups is to identify the most common surnames borne by immigrants to the mainland colonies or their descendants. This strategy ensures that the names chosen are representative of each nationality's population in the United States, even if their distribution differed from that in Europe.

[3] David Steven Cohen, "How Dutch Were the Dutch of New Netherland?" *New York History,* LXII (1981), 51-53; Israel Acrelius, *A History of New Sweden . . .* (1759), trans. William M. Reynolds (Historical Society of Pennsylvania, *Memoirs,* XI [Philadelphia, 1876]), 193, hereafter cited as Acrelius, *New Sweden.*

TABLE I
DISTRIBUTION OF DISTINCTIVE NATIONAL NAMES BY STATE

State	16 *Welsh Names*	71 *Scottish Names*	54 *Irish Names*	90 *German Names*	37 *Dutch Names*	11 *Swedish Names*	15 *French Names*
Me.	76	576	162	25			20
N.H.	111	779	178	2			9
Vt.	107	442	108	3	4		2
Mass.	461	1,406	331	25	13		55
R.I.	52	308	56	2			36
Conn.	256	737	169	19	4	1	15
N.Y.	381	1,901	443	618	1,180	2	162
N.J.	217	804	241	238	801	49	80
Pa.	544	4,503	1,034	3,485	134	72	49
Del.	91	304	131	27	14	49	10
Md.	299	1,347	698	508	19	16	69
Va.	913	3,296	945	389	66	21	118
N.C.	613	3,073	825	305	24	23	86
S.C.	323	1,951	421	175	7	11	67
Ga.	77	236	82	21	1	3	9
Ky.	61	559	151	51	14	3	9
Tenn.	79	556	139	66	14	5	5

Barker's attempts to ascertain the size of the German presence by extrapolating surname ratios suffered from a major shortcoming. Having no information on the frequency of family names in Germany, he chose for examination thirty-two common surnames from the Pennsylvania census. He then compared their distribution in other states relative to Pennsylvania to obtain percentages for the German element in the United States. He had no clear idea how many Germans actually lived in Pennsylvania, yet this information was absolutely essential for the accuracy of all figures based upon the incidence of the thirty-two names. Barker finally relied on the traditional wisdom about Pennsylvania's population—that it included English, Scotch-Irish, and Germans in roughly equal numbers—and fixed the German proportion at exactly 33.3 percent.[4] All his findings were thus predicated upon a conventional guess long current among Pennsylvania historians rather than upon any firm evidence of a quantitative nature.

A straightforward and accurate source for developing German surname ratios in eighteenth-century America is the compilation of foreign oaths of allegiance taken in Pennsylvania between 1727 and 1808. The surviving registers record the names of 29,758 German-speaking immigrants and indicate that these individuals brought 35,000 unidentified dependents with them, representing approximately two-thirds of all German arrivals in North America. These documents provide the best sample of names carried by the parent stock of the German-American population in 1790,

[4] "National Stocks," 291-294.

and enable a precise multiplier to be calculated for estimating the group's size. Ninety distinctive surnames included 12.5 percent of the 30,000 signers of the allegiance oath, equal to a numerical coefficient of 8.[5] By comparison, the thirty-two names chosen by Barker yield a constant of 23.4, a much higher figure that could bias the results toward overcounting this group.

Hansen based his estimates for Americans of Dutch, Swedish, and French descent upon an examination of the conditions under which each group settled in America and of the circumstances that influenced their subsequent growth and distribution. He then made projections of probable population increase based upon his knowledge of colonial demographic patterns. His conclusions were highly impressionistic and often amounted to little more than well-informed guesses. Hansen chose not to develop surname coefficients as Barker had done for Englishmen, Celts, and Germans, although he was aware of the very sources that would have enabled him to do so.

The Dutch population of the United States in 1790 was almost entirely descended from the seventeenth-century colonizers of New Netherland. The most direct way of devising a multiplier is to identify the Dutch names that were most common after emigration had ceased. Documentation from the late seventeenth and early eighteenth centuries has survived that enables the task to be completed for most Dutch inhabitants of New York. Between 1687 and 1714, a number of allegiance lists, tax rolls, and censuses were taken that enumerated 1,436 Dutch citizens of five counties and of New York City. The great majority of adult male Hollanders in the province were recorded; their total number can be determined quite accurately because in many counties with mixed populations, Dutch, English, Scots, and French were identified by nationality, depending upon which groups were in the minority. Fourteen percent of the Dutch on these lists had thirty-seven family names, yielding a coefficient of 7.3.[6]

[5] Ralph Beaver Strassburger, *Pennsylvania German Pioneers: A Publication of the Original Lists of Arrivals in the Port of Philadelphia from 1727 to 1808,* ed. William John Hinke, I (Norristown, Pa., 1934), xxxi. The names used are Albrecht (Allbright), Bachman, Baum (+Bohm), Baumgarten, Bentz, Brunner, Christ, Conrad, Diedrich, Diehl (Deal), Dietz, Doll, Eberhart, Eckhart, Ernst, Faust, Fink, Fritz, Funk, Gebhart, Geiger, Gerber, Gerhart, Graff, Grimm, Grubb, Hartman, Hahn, Hess, Hildebrand, Hoch (Hauk), Hoff, Hoffman, Houser, Huber (Hoover), Hummel, Jaeger (Yaeger), Jost (Yost), Kaufman, Keiser, Kesler, Klein, Kolb (Culp), Kramer, Kraft, Krebs, Kraus, Kuhn, Kuntz, Kurtz, Lentz, Link, Ludwig, Lutz, Metz, Metzger, Meyer, Moser, Ott, Rauch, Reinhart, Rudolph, Schafer, Schaub, Schneider (Snider), Schreiner, Schutz, Schultz, Schuman, Schuster, Schwaab, Switzer, Seibert, Seitz, Seifert, Seiler, Spengler, Spiess, Stahl, Stauffer (Stover), Stumpf, Stein, Steiner, Ulrich, Vogel, Wagner, Waltz, Weigandt, Ziegler, Zimmerman.

[6] The following documents were used: Allegiance List of 307 King's Co. Residents, Sept. 26-30, 1687, in E. B. O'Callaghan, ed., *The Documentary History of the State of New-York,* I (Albany, N.Y., 1849-51), 659-661; Allegiance List of 222 Ulster Co. Residents, Sept. 1, 1689, *ibid.,* 279-282; Census of 379 Albany Co.

Hansen's appraisal of the Swedes in America was just as unsystematic as his survey of the Dutch. He made rough estimates of the group's size along the Delaware River, sometimes examining tax lists or militia musters to supplement local histories. He did little more than assign arbitrary fractions of a percent in states south of Maryland to account for the small number there. Hansen tended to see Swedes everywhere and misjudged their number in many states.

A special census of Swedes living in the Delaware Valley in 1693 provides the necessary information to compile a roster of distinctive names and a corresponding multiplier. Nearly all Swedes a century later would have been the progeny of the 942 persons listed, since emigration from Scandinavia had ceased long before the taking of the 1693 census. Sixteen percent of the population bore eleven names, giving a coefficient of 6.3.[7]

Hansen found the task of determining the number of French-descended Americans in 1790 to be exceptionally difficult. Unlike the Dutch and the Swedes, few French colonists established separate communities whose growth and dispersion could be observed. Hansen made a state-by-state survey of known French migration and inspected the indexes to the 1790 census for evidence of Gallic names, but was unable to develop any precise means of measuring the group's size. He seems to have founded his final judgment on the assumption that the French in 1790 probably equaled one-half the Dutch, or 54,900, and then allocated this figure among the states as seemed most consistent with his examination of individual localities.[8]

No sources comparable to those for the Germans, Dutch, or Swedes exist that permit a coefficient for the French stock to be calculated from lists of colonists residing in British America. The majority of French

---

Residents, June 16, 1689, in Joel Munsell, ed., *The Annals of Albany,* IX (Albany, N.Y., 1858), 81-89; tax list of 49 Orange Co. residents, June 16, 1702, in O'Callaghan, ed., *History of N.-Y.,* I, 366; tax list of 818 New York householders about 1703, *ibid.,* 611-624; tax list of 87 Dutchess Co. residents, 1714, *ibid.,* 368-369. The King's and Albany lists identified non-Dutch residents, leaving a total of 617. All of those enumerated on the Ulster, Orange, and Dutchess lists were apparently Dutch, although a few may have been French refugees. Based upon Thomas J. Archdeacon's estimate that 56.3% of New York taxpayers in 1703 were Dutch, the assessment rolls would have included 461 Dutch (*New York City, 1664-1710: Conquest and Change* [Ithaca, N.Y., 1976], 46). The names used are Beekman, Bogart (Bogardus), Bradt, Covenhoven, Hageman, Hendrickson, Hogeboom, Houghtling, Kipp, Outhout, Quackenboss, Rapalye, Remsen, Schuyler, Swartwout, Ten Brock, Ten Eyck, Van Buren, Van Cleef, Vandeberg, Vandeveer, Vandewater, Van Dyke, Van Horn, Van Huten, Van Husen, Van Meter, Van Ness, Van Pelt, Van Shaike, Van Slyke, Van Vleet/Fleet, Van Wagonen, Van Wyck, Voorhees, Vosburgh, Wyckoff.

[7] It is printed in Acrelius, *New Sweden,* 193. The names used are Dalbo, Derickson (+Derek), Justison (+Justis), Matson, Paulson, Rambo, Sinnickson (Sinnex, Seneca), Stalcop, Stilley, Tussey (Tossa), Walraven.

[8] "National Stocks," 380-390.

settlers were Huguenots who fled to England and migrated overseas from there. Many of these Protestant refugees, and undoubtedly most who came to North America more than a generation after Louis XIV revoked the Edict of Nantes in 1685, were born in England. Consequently, examining the baptismal registers of Huguenot churches in England should provide a sample of French cognomens representative of the refugee community from which most emigrants originated. Seven percent of 26,932 children born in three French Reformed congregations from 1581 to 1750 held fifteen names, equal to a numerical constant of 14.[9]

Despite the obvious care taken by the McDonalds to isolate Irish, Scottish, and Welsh surnames, it is possible to refine their data to obtain more exact coefficients. Some modification is necessary because certain selected names are not truly representative of the ethnic groups under study and skew the results toward a high count of minority stocks if included. Using these names without any correction makes it impossible not to overestimate the proportions of Celtic peoples to some degree, because many English families also bear them. A relatively small number of English immigrants from these families residing in America by 1690 could have produced numerous descendants by 1790 because of the high birth rate in the colonies. The propensity for exaggerating minority stocks can be controlled if every effort is taken to ensure that all names chosen are as characteristically non-English as possible, and if compensation is made for any unavoidable overcount.

The McDonalds' Irish data suffer from the inclusion of three nondistinctive names that significantly inflate the results. If emigration from Ireland was representative of the island's native Gaelic stock, as must be assumed, then the total number of persons named Hayes, Hughes, and Johnston would compose 4.7 percent of all Irish-Americans in 1790 who bore the fifty-three surnames chosen by the McDonalds to compute the group's numerical constant. A figure substantially higher than 4.7 would indicate that the three names are unreliable and unduly exaggerate this community's size. In fact, the McDonalds found that these three names accounted for 26.5 percent of all individuals bearing the fifty-three surnames selected by them as distinctively Irish. Such a discrepancy is unjustifiably large, especially when one considers that holders of Ireland's three most common family names—Murphy, Kelly, and Sullivan—included only 15.6 percent of all Irish Americans with the same fifty-three names listed on the 1790 census.[10]

[9] The registers were from the French churches at Canterbury (1581-1728), Thorney in Cambridgeshire (1654-1727), and Threadneedle St. in London (1600-1750), in Huguenot Society of London, *Publications,* V (1891), 3-409, IX (1896), 34-228, XIII (1899), 70-281, XVII (1903), 1-103, XXIII (1916), 21-241. The distinctive names (which varied greatly in the original records) are Ballou, Bodine (Bowden), Cato (Caddo), Devine, Dubois, Dumont, Dupre, Durant, Fontaine, Lamar (Delmar), Larue, LeConte, LeFevre, LeMaster, Spain (DeSpain).

[10] Robert E. Matheson, *Special Report on Surnames in Ireland . . .* (Dublin, 1909 [orig. publ. 1894]), 7.

The excessive number of Johnstons, Hayeses, and Hugheses arises because all are common names elsewhere in the British Isles. Johnston is actually Scottish rather than Irish and is largely limited to Ulster. Since the McDonalds placed Boyle with the Scots for identical reasons, it is difficult to understand why they did not treat Johnston in the same manner.[11] Hughes is authentically Irish, but H. B. Guppy's research demonstrates that it is more common among the Welsh than the Irish and furthermore that it can be found in significant frequencies among the yeomen of one-quarter of England's counties.[12] Without compensating for the fact that many, if not most, of the persons named Hughes in America must have been of Welsh or English extraction, using the cognomen unduly increases the proportion of Irish. Hayes is not unusual in southern England, and it is common throughout Scotland, making its inclusion also undesirable.[13] Eliminating these three surnames and replacing them with the genuinely distinctive names of Cavanaugh, Riley, Sheehan, and Whelan leaves a multiplier of 5, which has the advantage of being smaller and so, as the McDonalds have pointed out, is slightly more accurate than the original constant of 5.2.

The degree to which the corrected Irish coefficient of 5 might also overstate that group's true proportion of the United States population can be estimated by making a rough calculation of how many individuals bore the fifty-four names in Ireland and then by comparing that number with the corresponding figure for England and Scotland. The Irish registrar-general concluded that 943,592 residents of Ireland had these names in 1891, while tabulations of surname frequencies among the rural population of Scottish and English counties listed by Guppy at approximately the same time (which exclude the Irish immigrant community there) indicate that the corresponding figure for those two was 16,590. The probable overcount of the Irish stock consequently is about 1.7 percent. Ireland's population had declined dramatically in relation to England's and Scotland's by 1891. Adjusting the overcount to reflect the distribution of Ireland's population within Great Britain a century earlier would reduce the estimated overstatement to much less than 1 percent, a figure so low that no correction of the revised coefficient is necessary.[14]

[11] *Ibid.*, 55. McDonald and McDonald, "Ethnic Origins," *WMQ*, 3d Ser., XXXVII (1980), 190n. Henry Brougham Guppy, *Homes of Family Names in Great Britain* (London, 1890), 590.

[12] Hughes occurs with a frequency of 42.6 per 10,000 in Wales and 31.5 per 10,000 in Ireland (Guppy, *Family Names*, 505-506; Matheson, *Surnames in Ireland*, 7).

[13] Guppy, *Family Names*, 498, 500, 589.

[14] The number of English and Scots bearing the 52 distinctive Irish names was obtained by multiplying the frequency of each name reported by Guppy for every county times the 1891 population and summing the results. Population figures were taken from B. R. Mitchell, *European Historical Statistics, 1750-1970* (London, 1975), 71-73. This procedure avoids including post-1845 migrants from Ireland because Guppy's work deals only with indigenous names borne by the yeomanry,

The problem of overcounting also exists for Scottish names. Guppy's work on British family nomenclature reveals that forty-three of the McDonalds' seventy-two names were found among the rural population of at least one English county in sufficient numbers to be considered indigenous to that locality. Were these occurrences restricted to the northern frontier area of Cumberland, Northumberland, Westmoreland, and Durham, they might cause minimal difficulty, but that is not the case. Thirty of the forty-three names were distributed south of the border shires, in every region of England. The great majority may be included because they are not especially numerous, but eight cause particular concern. Bell, Gibson, Scott, Shaw, Wallace (Wallis), Watson, and Williamson are common throughout England, while Kennedy is most numerous in Ireland outside of Ulster.[15] If emigrants from Scotland and Ulster carried the eight names to America in random fashion, as the methodology presumes, they would compose 10.7 percent of all individuals holding the seventy-two names according to the data cited by the McDonalds to calculate the Scottish multiplier.[16] Any significant difference from this figure could only result from inadvertently counting the descendants of Englishmen and Irishmen with the same names as Scots. Bearers of the eight questionable surnames constitute 18.5 percent of all the Scots listed in Table III of the McDonalds' article, a clear indication that Irish Kennedys and Englishmen holding the remaining cognomens have been unintentionally included. The simplest way to prevent the eight names from biasing the results is to exclude them.

Closely related to the problem of preventing nondistinctive names from exaggerating the numbers of Scotch-Irish and Scots is the McDonalds' insistence on excluding all patronymics prefixed by Mac and Mc. Their position has drawn serious criticism, since these are the very names *least* likely to skew the findings toward a high percentage of Scotch-Irish and Scots.[17] A better case exists for accepting their critics' suggestion than for rejecting it. The essential rationale for eliminating all Mac-preceded surnames was that English authorities discouraged their retention in the Highlands after the 1745 Jacobite uprising as a means of undermining the clan system.[18] Official efforts to suppress certain of these clan names must

---

not urban residents or inhabitants of manufacturing towns. Because of the extraordinary rise in the industrialized population of Lancashire and Yorkshire's West Riding from 1801 to 1891, their figures were excluded to prevent them from skewing the results. There were six Englishmen for every Irishman in 1891, a ratio more than three times greater than in 1750.

[15] The figures used by the McDonalds to determine the total number of Wallaces included the common English variant Wallis. See Bureau of the Census, *A Century of Population Growth: From the First Census of the United States to the Twelfth, 1790-1900* (Washington, D.C., 1909), 267; Guppy, *Family Names,* 455, 490, 547, 548, 566, 567, 572; and Matheson, *Surnames in Ireland,* 56.

[16] For the relative frequencies of these names, see "National Stocks," 213.

[17] Letter of Rowland Berthoff, *WMQ,* 3d Ser., XXXVII (1980), 701-702.

[18] McDonald and McDonald, "Ethnic Origins," *ibid.,* 191n.

have led some Scots to comply, but they seem to have failed in general since such patronymics continued to be widely held.

The most compelling reason to include names prefixed by Mac arises from the nature of the source used by the McDonalds to compute surname frequencies, James Stark's 1863 compendium of Scottish family names. Whatever the results of the British government's actions against the clan system, new nomenclatural patterns would have stabilized within a generation after 1745. Highlanders quickly adapted to new political and military realities, and by the Seven Years' War had successfully made the transition from fifth column to a reliable source of cannon fodder. Any decline in the use of names preceded by Mac developed within a short time (significantly prior to 1763, and so before the beginnings of heavy emigration from the Highlands) and would have been fully reflected in Stark's study of Scottish surname distributions, made over a century later. Stark's work, in other words, incorporated any shifts in naming customs that occurred after the 1745 uprising, making it unnecessary to adjust any further.

Perhaps the most important reason for the McDonalds' exclusion of such names (as they revealed later) was to ensure that any biases in the methodology would work *against* their Celtic Thesis. Counting the Macs would have increased the percentage of English in New England and of Scots in the Carolinas.[19] The McDonalds' integrity is admirable, but the best research strategy is to adopt the most reliable procedures for estimating all groups, without purposefully attempting to minimize or maximize any of them. Whenever overseas migration from Scotland was heaviest, the distribution of surnames most closely approximated the entire nation's, because a greater proportion of the Highland population left the country along with the Lowlanders, who usually predominated among Scottish emigrants. Consequently, it is legitimate to conclude that states where a large portion of the Scottish population carries names with Mac have a greater percentage of Scots than states where the portion is smaller. If the most accurate results support the Celtic Thesis, there is no reason not to present them. Precision depends on using the most distinctive surnames and the most characteristically Scottish cognomens are those with Mac.

It is doubtful in any case that English policies in Scotland significantly affected nomenclatural patterns in the colonies. Royal authorities did not extend prohibitions against clan names to loyal Highlanders, to Lowlanders, or to inhabitants of Ulster, from which most immigrants with Scottish names originated. The most common manner of anglicizing names after 1745 was simply to drop the Mac, which suggests a simple way of solving the problem identified by the McDonalds. M'Donald, M'Kenzie, M'Kay, M'Lean, M'Leod, M'Intosh, and M'Gregor can be added if derivations like Donald, Donnel, Cloud, Kay, Kenzie, and Gregor are counted as equivalent forms. This solution leaves a multiplier of 3.7, which has the

[19] Letter of Forrest and Ellen Shapiro McDonald, *ibid.*, 703.

advantage of being smaller and therefore theoretically a bit more accurate than the McDonald's original coefficient of 4.

The question of how greatly the resulting constant of 3.7 might exaggerate the Scottish stock in America can be investigated by comparing the probable number of persons with the seventy-one designated names in Scotland and Ireland with that in England. The Scottish registrar-general calculated in 1861 that 766,164 Scots bore these surnames. The best estimate of Irishmen holding the same names in 1861 would equal 336,342. The corresponding number of Englishmen in the same year was probably 112,519, a figure representing 9.3 percent of all persons with the distinctive "Scottish" names in Great Britain. The fact that the ratio of England's population to that of Ireland and Scotland in 1861 was much greater than it had been earlier, requires, however, that the overcount measurement be adjusted downward. Using statistics from Great Britain's first census in 1801 would reduce to 7 percent the English portion of the general population carrying the selected names. Relying on estimates made by twentieth-century demographers of the English, Scottish, and Irish populations in the 1750s would further decrease the projected overcount to approximately 6 percent. Given the range of different population statistics for the 1700s (figures published between 1955 and 1981 concerning England's 1751 population have varied by more than 6 percent), it is questionable whether a probable overcount of 6 percent should be corrected, since it would decrease the multiplier only by .2.[20] The coefficient for Scotch-Irish and Scots consequently may be left at 3.7.

The most frustrating problem in approximating the Scottish-descended population through surname frequencies is the inability to determine the proportion that migrated from Northern Ireland. Current research by R. J. Dickson and Ian C. C. Graham indicates that about 114,000 persons came to the American mainland from Ulster during the six decades after 1718, compared to perhaps 62,500 from Scotland.[21] These figures

[20] The procedure for estimating the English portion of the Scottish overcount is explained in n.14. The number of Scots bearing the 71 names is given in "National Stocks," 212-213. The number of Irishmen was determined by the method used by Matheson in *Surnames in Ireland,* multiplying the number of births for each name in 1890 by 44.8, the ratio of births to population, and summing the results. The resulting figure was then adjusted according to the ratio of 1890's population to that of 1861, and then according to population data available for the 1750s. The estimates used for the populations of England (5,772,000), Scotland (1,265,000), and Ireland (3,191,000) about 1750 were taken from E. A. Wrigley and R. S. Schofield, *The Population History of England, 1541-1871: A Reconstruction* (Cambridge, Mass., 1981), 577, Table A5.3; D. V. Glass, "Population and Population Movements in England and Wales, 1700 to 1850," in Glass and D.E.C. Eversley, *Population in History: Essays in Historical Demography* (London, 1965), 240; and Mitchell, *European Historical Statistics,* 34, 36. See the table cited above in Wrigley and Schofield for the range of estimates on England's 1751 population, such as Ohlin's (1955) of 5,450,000 and Lee and Schofield's (1981) of 5,790,000.

[21] The research of R. J. Dickson, *Ulster Emigration to Colonial America, 1718-1775* (London, 1966), 23, 34, 59, 64, indicates that the probable number of

indicate that the ancestors of approximately two-thirds of all Americans with Scottish names in 1790 had migrated from Northern Ireland rather than Scotland. Dividing the Scottish stock by this ratio can provide a rough estimate of the number of Scots and Scotch-Irish. This procedure is admittedly arbitrary, but it is the only solution capable of distinguishing between the relative contributions of each group at the state and national levels.

The McDonalds' lists of representative names for Scots and Irish have required relatively little adjustment to ensure their reliability, but the Welsh present a major problem. The McDonalds estimated the Welsh stock by using the twelve names of Edwards, Evans, James, Jones, Lewis, Morgan, Morris, Phillips, Price, Roberts, Thomas, and Williams. They should have taken into consideration that Guppy, the authority on English names, demonstrated that these "Welsh" cognomens were common throughout England, except in the extreme north, and were very prominent in the West Country, not simply the Celtic fringe. Guppy identified most of the twelve as native to more than half of England's counties.[22] If the numbers of persons holding these surnames in 1801 (when Great Britain conducted its first census) are estimated according to the frequencies and distribution given for each county's rural population by Guppy, then only 61 percent would have been residents of Wales or the Welsh-speaking shire of Monmouth rather than of England. Overrepresentation of Englishmen on such a scale requires either a drastic reduction in the original multiplier from 2.5 to 1.5 or the development of a new coefficient based on a more representative group of family names.[23] It is worth trying the latter course to see if a more accurate testing procedure can be devised.

Approximately 15 percent of the Welsh people bear the sixteen names of Bebb, Breese, Griffith, Howell, Humphreys, Jenkins, Lloyd, Morgan, Owen, Pritchard, Pugh, Rees, Rowlands, Tudor, Vaughan, and Wynne, equal to a multiplier of 6.5. These names are much less common among Englishmen than the McDonalds' twelve. Guppy's research found that only ten of the sixteen names occurred among the yeomanry of as many as

---

Ulster-Scot emigrants ranged between 102,000 and 125,000, leaving 114,000 as a reasonable middle figure. According to Ian Charles Cargill Graham, *Colonists from Scotland: Emigration to North America, 1707-1783* (Ithaca, N.Y., 1956), 185-188, emigration was "sporadic" before 1763 but reached 24,500 in the next 12 years. The average migration of 2,000 persons per year was the largest loss experienced in the period after 1707 in Graham's estimation, indicating that only a few hundred Scots left for America in most years before 1763. The number of emigrants probably ranged between 50,000 and 75,000, with 62,500 the most likely figure.

[22] Guppy, *Family Names,* 481, 483, 508, 509, 515, 525, 535, 538, 543, 559, 572.

[23] The number of English and Welsh bearing the 12 distinctive names was determined by multiplying the frequency of each surname required for every county by Guppy times the 1801 population and summing the results. See Mitchell, *European Historical Statistics,* 71-72, for figures used.

six English shires in ratios exceeding six per ten thousand, and none could be considered indigenous to more than eight of the forty English counties in 1890. The 1801 census indicates that there were most likely 88,050 persons then holding the sixteen names in Wales and Monmouth, compared to about 28,346 in England.[24] The probable level of overcounting, 24 percent, requires a reduction in the multiplier to 4.9. This study will use the sixteen names listed above for calculating the Welsh stock because they are found less frequently in England than the twelve chosen by the McDonalds and so are less likely to exaggerate this group's numbers in America, but still are capable of producing a sufficiently small coefficient to guard against biasing the results toward an unduly large number of Welsh.

Having revised the Irish, Scottish, and Welsh multipliers, it remains to ask how precise is any measure of nationality based upon European nomenclatural frequencies. One way of testing the method's accuracy is to see whether two different lists of distinctive names for a particular nationality can project estimates that differ only slightly. If the method is indeed reliable, then any two groups of characteristic family names should produce close results once their respective coefficients have been adjusted to compensate for overcounting. An indication of how accurately the method measures European descent can be gained by comparing the percentage of Welsh found by using the McDonalds' twelve cognomens and their corrected constant of 1.5 with the corresponding figure given by the sixteen names listed above and their revised multiplier of 4.9.

The first twelve names recur 15,359 times on the census rolls and tax lists of the seventeen states covered by the first United States census; when increased by 1.5, the number of Welsh heads of families rises to 23,033, equivalent to 4.43 percent of all white Americans in 1790. The sixteen other names occur 4,661 times on the same sources, indicating that the total number of Welshmen heading households equaled 22,839, 4.38 percent of the European-descended population. The difference between the projections is a remarkably small 194 individuals, less than 1 percent of either's total. If precautions are taken to prevent overcounting, it would seem that the method employed here gives a reliable approximation of the degree to which different European nationalities contributed to American demographic growth.

The precision with which the Welsh stock was estimated using two different sets of distinctive surnames suggests that the margin of error for other minority stocks is also low but gives no indication how low it might be. If the range of inaccuracy were six times that suggested by testing the Welsh coefficients, the degree of error would be about 5 percent. An average variation of 5 percent would mean that the national figures given for each group identified in Table II could be subject to the following level of uncertainty: French, ±.1; Dutch, ±.2; Irish, ±.3, Scots, ±.3; Welsh, ±.2; Germans, ±.4; Scotch-Irish, ±.5. The cumulative error in

[24] See n. 20 for the procedures used in estimating the overcount.

TABLE II

REVISED CLASSIFICATION OF THE WHITE POPULATION BY NATIONAL OR LINGUISTIC STOCKS, 1790

State	*English* %	*Welsh* %	*Scotch-Irish* %	*Scottish* %	*Irish* %	*German* %	*Dutch* %	*French* %	*Swedish* %	*Total* %
Me.	77.6	2.2	8.4	4.2	4.8	1.2		1.6		100
N. H.	81.4	2.3	8.0	4.0	3.7	.1		.5		100
Vt.	81.4	3.5	7.3	3.6	3.6	.2	.2	.2		100
Mass.	84.4	3.5	5.3	2.7	2.5	.3	.1	1.2		100
R.I.	79.9	2.3	7.0	3.5	2.6	.1		4.6		100
Conn.	87.1	3.1	4.5	2.2	2.1	.4	.1	.5		100
N.Y.	50.3	3.4	8.7	4.3	4.1	9.1	15.9	4.2		100
N.J.	50.6	3.6	6.8	3.4	4.1	6.5	20.1	3.8	1.1	100
Pa.	25.8	3.6	15.1	7.6	7.1	38.0	1.3	.9	.6	100
Del.	63.3	5.5	9.2	4.6	8.0	2.6	1.3	1.7	3.8	100
Md.	52.5	4.6	10.4	5.2	10.9	12.7	.4	3.0	.3	100
Va.	61.3	6.5	11.7	5.9	6.8	4.5	.7	2.4	.2	100
N.C.	53.2	6.2	15.8	7.9	8.6	5.1	.4	2.5	.3	100
S.C.	47.6	6.2	18.9	9.4	8.2	5.5	.2	3.7	.3	100
Ga.	58.6	7.9	12.2	6.1	8.6	3.5	.1	2.6	.4	100
Ky.	54.8	3.6	16.5	8.3	9.0	4.9	1.2	1.5	.2	100
Tenn.	50.6	4.8	17.8	8.9	8.7	6.6	1.3	.9	.4	100
U.S.	59.7	4.3	10.5	5.3	5.8	8.9	3.1	2.1	.3	100

understating or inflating the English presence in the total population could be as high as two percentage points if all variation occurred in the same direction; this eventuality would nevertheless be unlikely since some errors would inevitably offset others, leaving a probable range of inaccuracy of 1. The obvious limitations of the data make some inaccuracy unavoidable; but even if the degree of error were 10 percent, none of the minority stocks would be understated or overestimated by more than one percentage point at the national level, and the English estimate would probably be within two or three percentage points of the true figure.

The most uncertain statistics on national origins have always concerned the states whose 1790 census manuscripts perished when the British burned Washington in 1814: New Jersey, Delaware, Georgia, Kentucky, Tennessee, and Virginia. Barker and Hansen obtained estimates for all but Virginia by interpolating from the ethnic mix found in neighboring states. Fortunately, alternative sources exist which enable reasonably reliable projections to be made for each. Delaware's census rolls from 1800 have survived; they provide the best substitute for the 1790 schedules, since the state's population underwent no basic changes in the preceding decade aside from natural increase.[25] Tax lists can be used to replace missing census records for the remaining states. The initial assessment rolls from each township in New Jersey during the 1770s or 1780s identify 29,162 individuals, only 600 fewer than the number of household heads estimated to have resided there in 1790.[26] Surviving lists of Kentucky taxpayers from 1790 exist that name 8,352 individuals, almost three-quarters of the number of heads of families believed to have been residing there at the time of the first census.[27] Assessment records for Tennessee and Georgia are not as full as those for New Jersey and Kentucky. A collection of the oldest tax digests of Georgia taken between 1792 and 1812 permits 4,780 ratables to be studied, approximately one-half the number of men who would have been recorded by the first census.[28] Tax rolls dated from 1787 to 1812 enumerate 7,963 residents of East Tennessee, the region where the great majority of the territory's settlers lived in 1790, from which an estimate of ancestral origins can be made.[29] Previous calculations of Virginia's population have relied on a collection of tax returns published by the Census Bureau from half of the state's counties in the 1780s. A

[25] Gerald and Doris Ollar Maddux, *1800 Census of Delaware* (Baltimore, 1964). The census lists 8,152 heads of household.

[26] Kenn Stryker-Rodda, *Revolutionary Census of New Jersey . . .* (Cottonport, La., 1972). Census Bureau, *Century of Population Growth,* 96.

[27] [Charles Brunk Heinemann and Gaius Marcus Brumbaugh], *"First Census" of Kentucky, 1790* (Washington, D.C., 1940). Census Bureau, *Century of Population Growth,* 96.

[28] Ruth Blair, ed., *Some Early Tax Digests of Georgia* (Atlanta, Ga., 1926). Census Bureau, *Century of Population Growth,* 96. Camden, Glynn, Wilkes, Montgomery, Chatham, Hancock, and Warren counties were used.

[29] Pollyanna Creekmore, *Early East Tennessee Taxpayers* (Easley, S.C., 1980). Campbell Co. (1818-1833) and Cocke Co. (1839-1840) lists were not used.

supplementary listing of ratables in all the remaining counties was later printed and should be used in conjunction with the first. Between them, these two editions include 69,223 Virginians, 91 percent of the heads of households whose names were lost when the missing population returns burned in 1814, and enable statistics for the state to be projected with a high degree of certitude.[30] The total number of entries for the distinctive names of each national stock or linguistic group is displayed in Table I.

Table II shows the revised estimates of national stocks within the white population as of 1790. More precise and systematic research has modified Barker and Hansen's conclusions in several important respects. The distributions of ethnic groups in New Jersey, Delaware, Georgia, Kentucky, and Tennessee show numerous changes—not surprisingly, given the unsystematic nature of Barker's and Hansen's calculations. Separate data for Kentucky and Tennessee show these states' populations not to have been entirely alike, as had been earlier assumed. For the nation as a whole, the proportions of Americans descended from minor stocks is represented far more accurately than Hansen was able to do in his impressionistic survey. The percentage of Germans rises or falls significantly in several states, particularly in Pennsylvania where it increases from 33 to 38 percent. The numbers of English, Scots, and Irish fall between those of Barker and the McDonalds, but are closer to the latters'. With the addition of new statistics for the five states not analyzed by the McDonalds, it now becomes possible to update figures on the size of every European stock at the national level for the first time in more than a half century.

The data in Table II have undeniable limitations. The present study considers national stocks as defined by language or by traditional civil boundaries in Great Britain. The figures describe only ancestral origins, and solely in the male line. The percentages admittedly oversimplify the ethnic backgrounds of virtually every group's members by failing to distinguish between such diverse regions as the Scottish Highlands and Lowlands, the German Palatinate and Switzerland, and the several distinctive regions of England. Classifying national stocks in this manner simply recognizes that, although national descent can be estimated with reasonable precision, the extraordinarily high rate of internal migration within Great Britain and other European countries during the early modern period prevents any methodology premised upon nomenclatural frequencies from identifying regional origins with any assurance of accuracy.

A further limitation of the study is its inability to account for the effects of acculturation and intermarriage. Both of these factors must have blurred the boundaries between national stocks by 1790 and in some cases obliterated them. It is extremely doubtful whether descendants of Scot-

[30] Bureau of the Census, *Heads of Families at the First Census of the United States Taken in the Year 1790: Virginia* (Washington, D.C., 1908). Augusta B. Fothergill and John Mark Naugle, *Virginia Tax Payers, 1782-87: Other Than Those Published by the United States Census Bureau* ([Richmond, Va.], 1940).

tish, Irish, and Welsh immigrants to New England or the South during the seventeenth century retained any Celtic characteristics that would have distinguished them from their Anglo-American neighbors after three, four, or perhaps five generations of close association and intermarriage. The Welsh, for example, composed 6 percent of all persons landing in Virginia from 1623 to 1666, nearly equal to their portion of the state's white population in 1790. The descendants of these immigrants (who were probably a large majority of those identified as Welsh in 1790) could not have maintained any substantial degree of cultural distinctiveness after having been scattered randomly among the Anglo-Virginian majority for more than a century without the sort of religious or communal institutions that could reinforce ethnic customs.[31] Ethnic exclusiveness seems to have been most characteristic of the Middle Atlantic region and the southern backcountry, where settlement patterns minimized contact between groups and the population derived from more recent immigrants.

Strictly speaking, figures on national descent measure only the extent to which several waves of European immigration had contributed to overall population growth in the United States by 1790. The major deficiency with the results reported here is that they probably underestimate the English by an unknown percentage, although, it is hoped, eliminating nondistinctive Irish and Scottish names and compensating for the high frequency of Welsh names within England have kept the problem of overcounting minority stocks to a minimum. Despite these difficulties, the revised data are as accurate as existing sources and research into nomenclatural frequencies permit at this time. The improved statistics describe the distribution of national stocks along the eastern seaboard with much more exactness than before, and provide better information for assessing the influence of different European groups upon demographic growth, frontier expansion, and the evolution of regional folkways.

[31] George Cabell Greer, *Early Virginia Immigrants, 1623-1666* (Baltimore, 1960 [orig. publ. Richmond, Va., 1912]), claims to list nearly 25,000 names in alphabetical order, but the actual total is approximately 18,000. Besides the Welsh, 3.7% of pre-1667 immigrants held Scottish names and 1.1% had Irish names, indicating that by 1690, after which Virginia's population grew primarily by natural increase, about 11% of the colony was of Celtic extraction. This group's progeny probably composed the same proportion of the state's population in 1790, but must have been completely assimilated into a society dominated by Anglo-Virginians. Consequently, to conclude from Table II that all 31% of Virginians identifed as Welsh, Scots, or Irish shared a common Celtic identity is unwarranted, since at least one-third could not have differed in any significant way from the English-descended majority, and most were undoubtedly of partial English descent themselves.